

dbGaP Study Release Notes



Release Notes for NHLBI TOPMed WGS CHIRAH, phs001605.v2.p1

"NHLBI TOPMed: Chicago Initiative to Raise Asthma Health Equity (CHIRAH)"

For any questions or comments, please contact: dbgap-help@ncbi.nlm.nih.gov.

May 6, 2020 Version 1 Data set release date
August 11, 2021 Version 2 Data set release date

2021-08-11

Version 2 Data set release for NHLBI TOPMed WGS CHIRAH now available

This release includes updated phenotype tables, the addition of Freeze 9 whole genome sequences (WGS), and corresponding VCFs. Please refer to the latest study configuration report for a detailed description of each download component.

Consent group 2 (c2): Disease-Specific (Asthma, IRB, COL) (DS-ASTHMA-IRB-COL)

Data Type	subjects	samples
Phenotype	292	296
Seq_DNA_SNP_CNV (VCFs)	276	276
WGS	276	276

For a description of SAMPLE_USE terms, please see:

<https://www.ncbi.nlm.nih.gov/projects/gap/submission/GetSampleUseTypes.cgi>

Study and Phenotype Updates

1. New Study Accession

NHLBI TOPMed WGS CHIRAH version 1 phs001605.v1.p1 has been updated to version 2. The dbGaP accession for the current data is **phs001605.v2.p1**. The participant number (p#) has not changed in version 2. No new subjects have been added to this study.

2. Updated Datasets (n=3 datasets)

pht	version	Dataset Name
10079	2	TOPMed_WGS_CHIRAH_Subject
10080	2	TOPMed_WGS_CHIRAH_Sample
10082	2	TOPMed_WGS_CHIRAH_Sample_Attributes

3. New Variable (n=1 variable)

pht	pht version	Dataset Name	phv	Variable Name
10079	2	TOPMed_WGS_CHIRAH_Subject	492525	SEX

4. Retired Variable (n=1 variable)

pht	Dataset Name	phv	version	Variable Name
10079	TOPMed_WGS_CHIRAH_Subject	429988	1	AFFECTION_STATUS

Molecular Data Updates

1. See download components 'sample-info' for manifest of genotyped samples and files.
2. Genotype calls are available in matrix format ('genotype-calls-vcf') as VCF file(s) with samples of same data sharing consent.
3. Quality control data are in 'genotype-qc' tars.

phg	freeze	sample_cnt	subject_cnt
phg001396.v1	8	267	267
phg001618.v1	9	276	276

Authorized Access (Individual Level Data)

Individual level data are available for download through the dbGaP Authorized Access System upon approval of the Data Access Request (DAR):

- <https://dbgap.ncbi.nlm.nih.gov/aa/wga.cgi?login=&page=login>

Public FTP site (Summary Level Data Only)

All data tables, data dictionaries, and documents will be housed under one directory for ease of downloading. The data_dict filenames have an added study version number (phs#.v#) and deleted participant set number (p#) from the table accession (pht#.v#). The var_report filenames have an added study version number (phs#.v#). In the var_report files, variables contain version numbers (phv#.v#) and summaries were created for each consent group (c#). These FTP files are available at:

- <https://ftp.ncbi.nlm.nih.gov/dbgap/studies/phs001605/phs001605.v2.p1>

2020-05-06

Version 1 Data set release for NHLBI TOPMed WGS CHIRAH now available

This release includes the addition of Freeze 8 whole genome sequences (WGS) brokered through the Sequence Read Archive (SRA), and VCFs derived from WGS. Please refer to the latest study configuration report for a detailed description of each download component.

Consent group 2 (c2): Disease-Specific (Asthma, IRB, COL) (DS-ASTHMA-IRB-COL)

Data Type	subjects	samples
Phenotype	292	292
Seq_DNA_SNP_CNV (VCFs)	267	267
WGS*	267	267

*These data are brokered through the Sequence Read Archive (SRA). Please see Authorized Access instructions below.

For a description of non-SRA SAMPLE_USE terms, please see:

<https://www.ncbi.nlm.nih.gov/projects/gap/submission/GetSampleUseTypes.cgi>

Molecular Data

dbGaP Study Release Notes



1. See download components phg001396.v1.TOPMed_WGS_CHIRAH.sample-info.MULTI.tar for manifest of genotyped samples and files.
2. Genotype calls are available in matrix format ('genotype-calls-vcf') as VCF file(s) with samples of same data sharing consent.
3. Quality control data are in phg001396.v1.TOPMed_WGS_CHIRAH.genotype-qc.MULTI.tar.

Authorized Access (Individual Level Data and SRA Data)

Individual level data and Sequence Read Archive (SRA) data are available for download through the dbGaP Authorized Access System upon approval of the Data Access Request (DAR):

- <https://dbgap.ncbi.nlm.nih.gov/aa/wga.cgi?login=&page=login>

Public FTP site (Summary Level Data Only)

All data tables, data dictionaries, and documents will be housed under one directory for ease of downloading. The data_dict filenames have an added study version number (phs#.v#) and deleted participant set number (p#) from the table accession (pht#.v#). The var_report filenames have an added study version number (phs#.v#). In the var_report files, variables contain version numbers (phv#.v#) and summaries were created for each consent group (c#). These FTP files are available at:

- <https://ftp.ncbi.nlm.nih.gov/dbgap/studies/phs001605/phs001605.v1.p1>